



# 图像分类任务 (2)

叶山 中国地质大学 (北京)

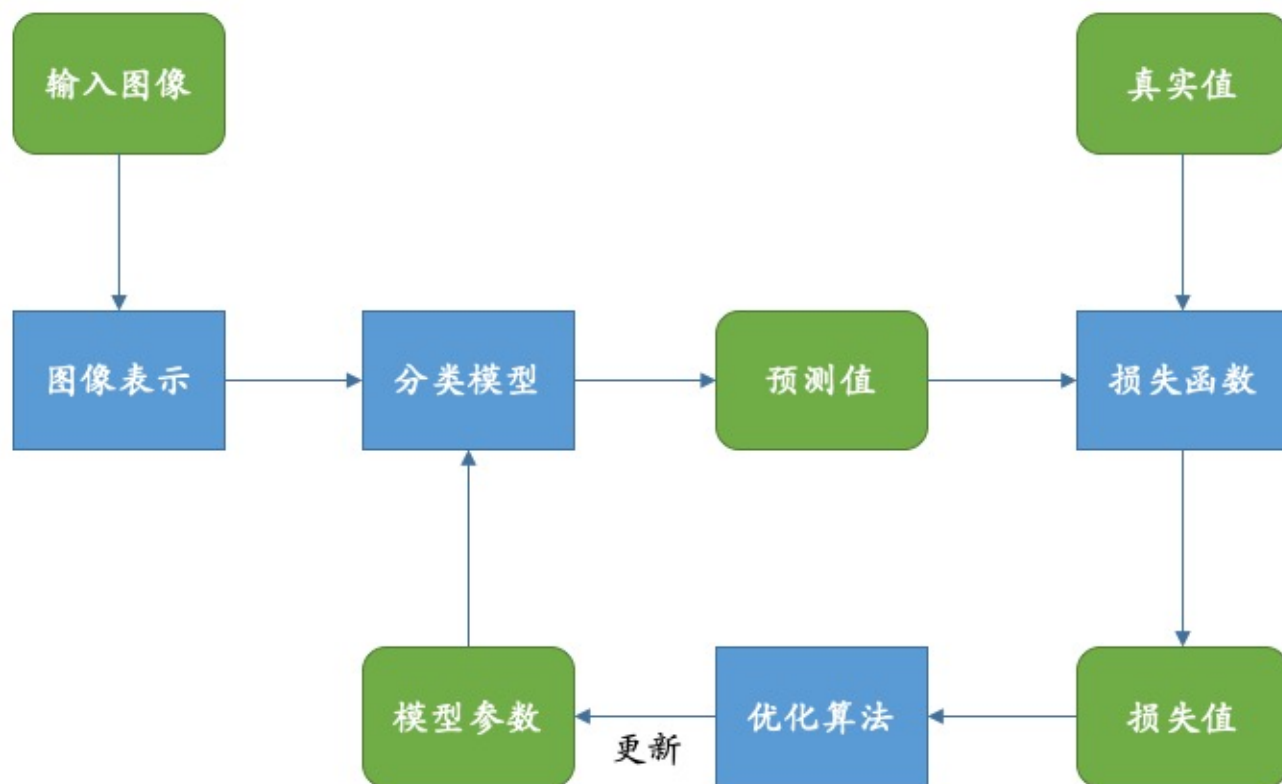
yes@cugb.edu.cn

# 数据驱动的图片分类

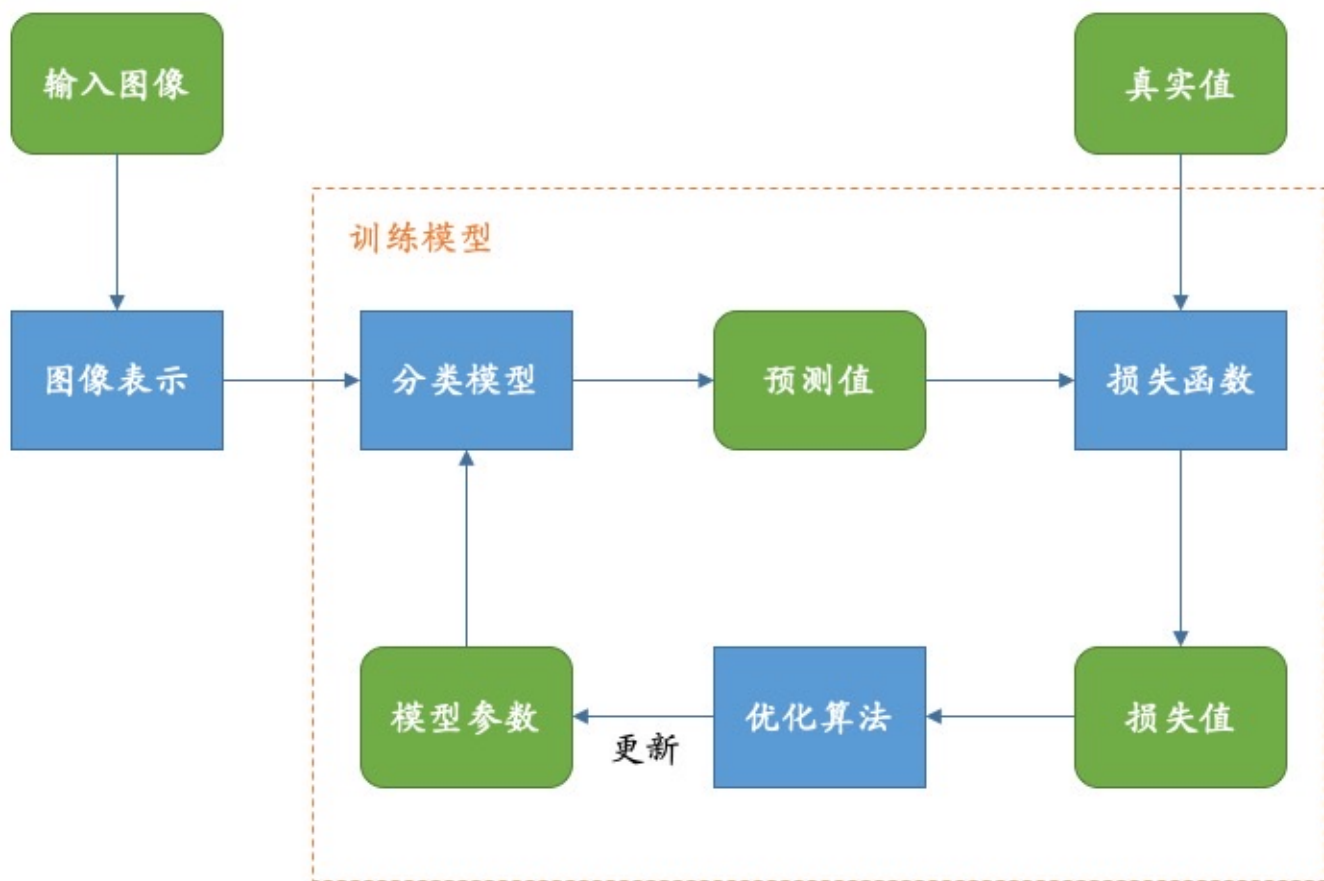
# 数据驱动的分类步骤



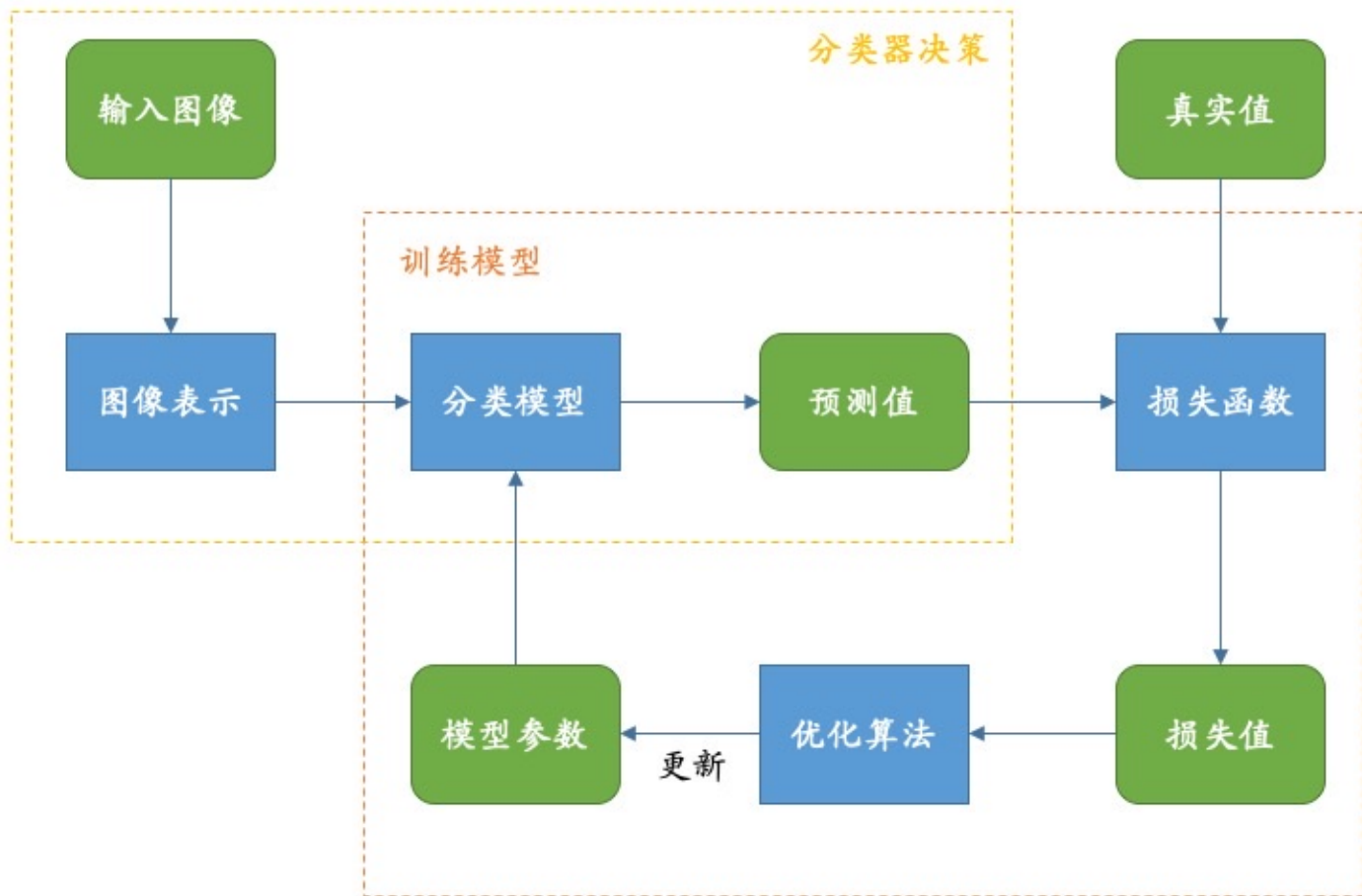
# 分类器的设计



# 分类器的设计



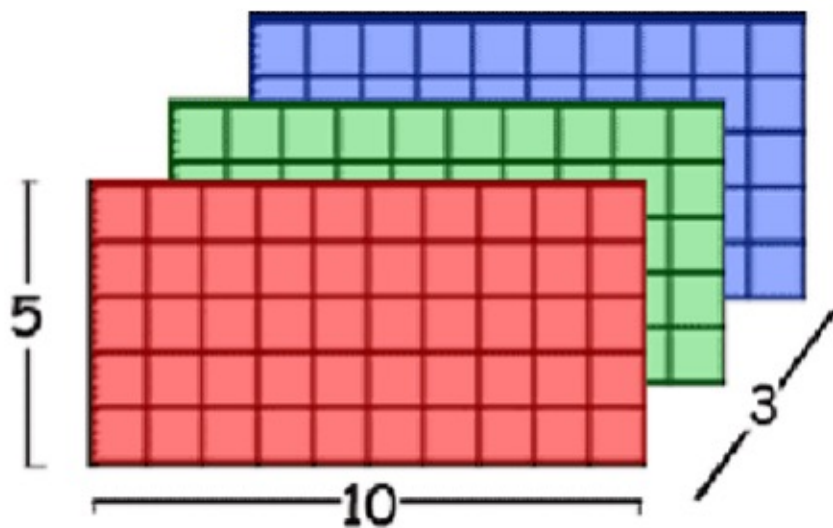
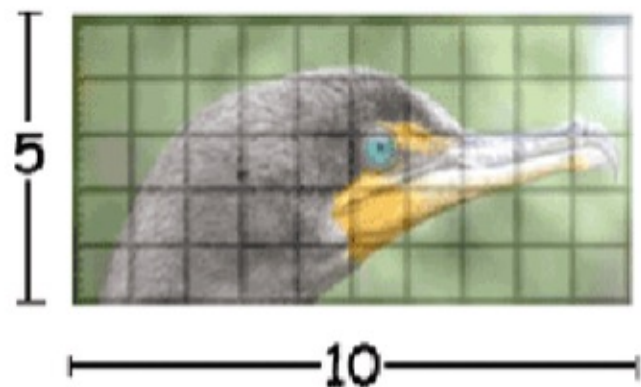
# 分类器的设计



# 图像表示

- 像素表示
- 全局特征表示（比如GIST）
- 局部特征表示（比如bag of feature模型）

# 像素表示



最简单、最基本的图像表示。



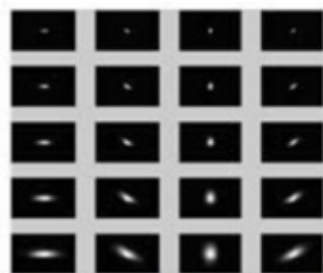
# 全局特征表示

将图像划分为若干个小区域（通常是4\*4共16个单元格）。

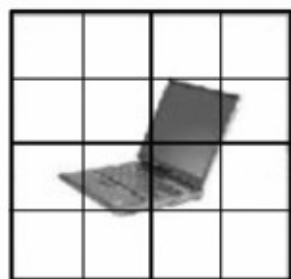
使用一系列不同的Gabor滤波器组成的滤波器组，对图像进行滤波。

计算每个单元格内的部分对每个滤波器的响应值，并计算平均值。

最终得到的GIST描述符的大小为 $4*4*N$ ，其中N是滤波器组里的滤波器个数。

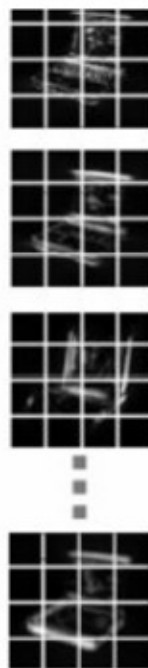


Gabor滤波器组  
5个频率4个方向



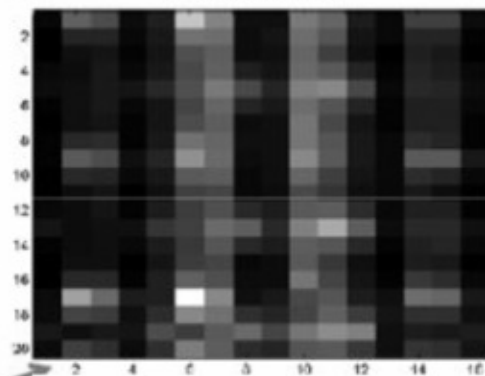
原图  
划分为16个网格

Gabor滤波器组进行滤波，得到20个结果



计算每一个网格内的滤波结果的均值

GIST全局表示



将均值排列成矩阵  
每一行对应一个滤波器  
每一列对应一个网格

# 局部特征表示



# 局部特征表示

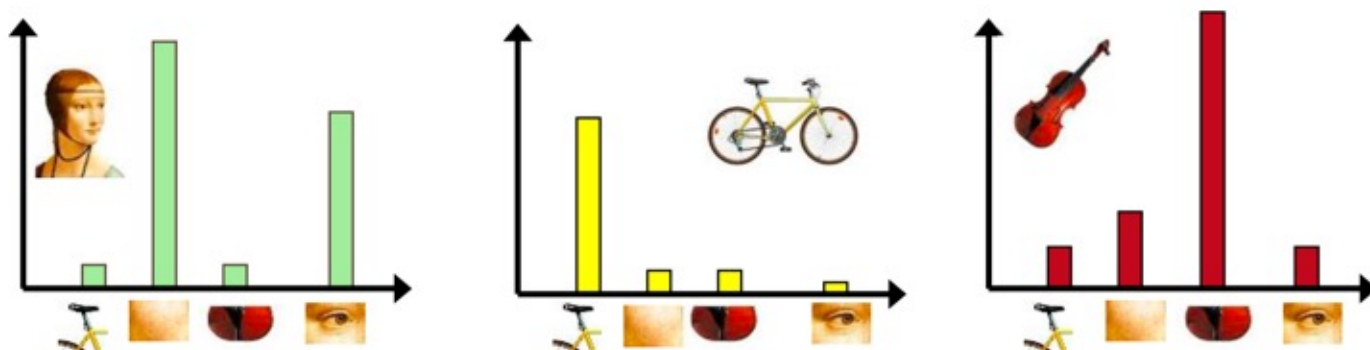
2007-01-23: State of the Union Address George W. Bush (2001-)

1962-10-22: Soviet Missiles in Cuba John F. Kennedy (1961-63)

1941-12-08: Request for a Declaration of War Franklin D. Roosevelt (1933-45)

abandoning acknowledge aggression aggressors airplanes armaments **armed** army assault assembly authorizations bombing  
 britain british cheerfully claiming constitution curtail december defeats defending delays democratic dictators disclose  
 economic empire endanger **facts** false forgotten fortunes france **freedom** fulfilled fullness fundamental gangsters  
 german germany god guam harbor hawaii hemisphere hint hitler hostilities immune improving indies innumerable  
**japanese** labor metals midst midway navy nazis obligation offensive  
 officially pacific partianship patriotism pearl peril perpetrated perpetual philippine preservation privilege reject  
 repaired resisting retain revealing rumors seas soldiers speaks speedy stamina **strength** sunday sunk supremacy tanks taxes  
 treachery true tyranny undertaken victory **war** wartime washington

Bag of words 词袋模型  
早期自然语言处理



Bag of features  
图像描述



# 图像表示

抓住重点维度，去掉和问题无关的维度。

# 分类器

近邻分类器  
(KNN)

贝叶斯分类器

线性分类器

支持向量机  
(SVM)

随机森林 (RF)

人造神经网络  
(ANN)

自适应增强  
(AdaBoost)

给输入图像打上分类标签，输出预测值

# 损失函数

0-1损失

多类支持向量机  
损失

交叉熵损失

L1损失

(平均绝对误差)

L2损失

(均方误差)

评估预测值和真实值之间的差距

# 优化算法

## 一阶算法

- 梯度下降
- 随机梯度下降
- 小批量随机梯度下降

## 二阶算法

- 牛顿法
- 拟牛顿法 (如DFP)
- BFGS
- L-BFGS

指导模型在学习过程中的参数迭代更新

# 训练过程

数据集划分

数据预处理

数据增强

欠拟合和过拟合

超参数和调参

模型集成



# Pop Quiz

1. Gaussian滤波器的频带宽度由什么决定？

A. 标准差 | B. 均值 | C. 极差 | D. 中位数

2. 以下哪个边缘检测器的边缘定位最不精准？

A. Sobel | B. Prewitt | C. Laplacian | D. Canny

3. 以下哪个方法被直接用于平滑图像、去除噪声？

A. 高通滤波 | B. 低通滤波 | C. 快速傅里叶变换 | D. 小波变换

4. 跨越语义隔阂指的是建立从\_\_\_\_到\_\_\_\_的映射？

A. 像素到语义 | B. 语义到像素 | C. 维度到语义 | D. 语义到维度

5. 发明时间最早、波形最简单的小波是？

A. Morlet | B. Meyer | C. Haar | D. Daubechies

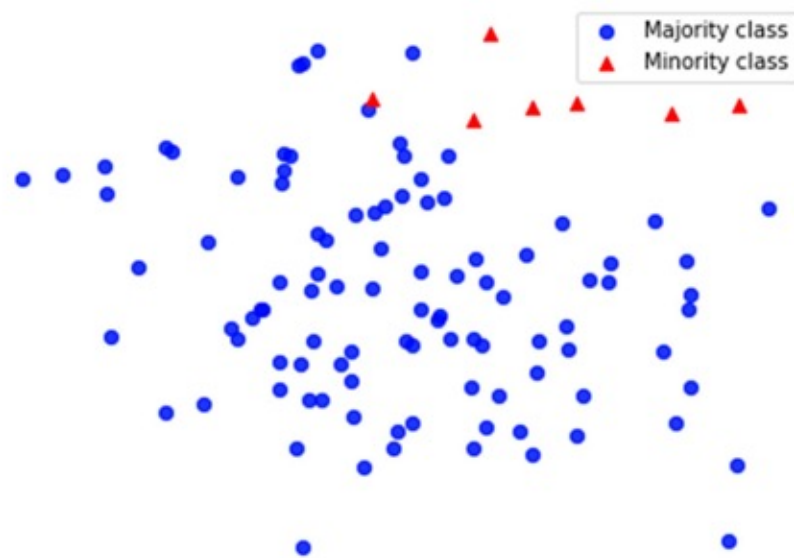


# 图像分类任务的评价指标

# 正确率Accuracy

分类正确的样本数占全部样本数之比。

- 优点：计算简单、易于理解
- 缺点：对于不平衡的数据集，因为majority class的影响，正确率的评估效果可能不好
- 需要额外指标：Precision（精准度），Recall（召回率）和F1-score



# 混淆矩阵 Confusion matrix

		分类结果	
		N 阴性	P 阳性
真实值	N 阴性	TN 真阴性	FP 假阳性 Type 1 Error
	P 阳性	FN 假阴性 Type 2 Error	TP 真阳性

True Positive (TP): 把阳性样本正确地分为阳性类。

True Negative (TN): 把阴性样本正确地分为阴性类。

False Positive (FP): 把阴性样本错误地分为阳性类。

False Negative (FN): 把阳性样本错误地分为阴性类。

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

# 精准度Precision和召回率Recall

		分类结果	
		N 阴性	P 阳性
真实值	N 阴性	TN 真阴性	FP 假阳性 Type 1 Error
	P 阳性	FN 假阴性 Type 2 Error	TP 真阳性

Precision: 在所有被分类为阳性的数据中，真实的阳性数据占多少？

Recall: 在所有真实值为阳性的数据中，有大比例被正确分类为阳性？

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

# F1 score

		分类结果	
		N 阴性	P 阳性
真实值	N 阴性	TN 真阴性	FP 假阳性 Type 1 Error
	P 阳性	FN 假阴性 Type 2 Error	TP 真阳性

综合考虑Precision和Recall。

$$F1 = (2 * Precision * Recall) / (Precision + Recall)$$

# 多于两类的情况

		Predicted		
		Cat	Dog	Pig
Actual	Cat	40	20	10
	Dog	35	85	40
	Pig	0	10	20

每个类别分别计算：

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

Macro平均：求每个类别的Precision、Recall、F1 score的均值。

Weighted平均：根据每个类别的样本数，每个类别的Precision、Recall、F1 score的加权均值。

		Predicted		
		Cat	Dog	Pig
Actual	Cat			FP
	Dog			FP
	Pig	FN	FN	TP

		TP	FN	FP
		Cat	40	30
Dog	85	75	30	
Pig	20	10	50	

